IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant: Baruch, et al.
Serial No.: 10/005,314
Filed: December 7, 2001
Title: VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES

Group Art Unit: 2655
Examiner: B. Albertall

REQUEST TO CORRECT INVENTORSHIP UNDER 37 CFR § 1.48(a)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

It has been found that the above-referenced patent application, through error and without deceptive intent, currently improperly sets forth the inventorship of the invention claimed therein. Accordingly, Applicant hereby respectfully requests that the above-referenced application be corrected in compliance with 37 CFR § 1.48(a), and that the following person be added as inventor:

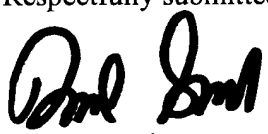
Eran Aharonson, a citizen of Israel, residing at 27 Hasaifan St., Ramat Hasharon, Israel.

Please find included herewith:

- A statement from each person being added as an inventor that the error in inventorship occurred without deceptive intention on his or her part;
- An oath or declaration by the actual inventor or inventors;
- A Revocation/Power of Attorney by the assignee;
- The written consent of the assignee;
- A statement under 37 CFR § 3.73(b).

The processing fee set forth in § 1.17(i) is believed to be due in connection with the filing of this request. The United States Patent and Trademark Office is therefore hereby authorized to charge Deposit Account No. 501380 in the amount of One Hundred and Thirty Dollars ($130.00) and any other fees that may be required for entry of this request.

Respectfully submitted,



Daniel J. Swirsky
Representative for Applicant
Registration No. 45,148

DANIEL J. SWIRSKY
ALPHAPATENT ASSOCIATES LTD.
55 REUVEN
P.O.B. 2345
BEIT SHEMESH, 99544
ISRAEL
TEL. (US) 516-620-4572
FAX. (US) 206-374-6672

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant: Baruch, et al.
Serial No.: 10/005,314
Filed: December 7, 2001
Title: VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES
Group Art Unit: 2655
Examiner: B. Albertall

STATEMENT BY INVENTOR UNDER 37 CFR § 1.48(a)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

It has been found that the above-referenced patent application currently improperly sets forth the inventorship of the invention claimed therein. I hereby respectfully request that the above-referenced application be corrected in compliance with 37 CFR § 1.48(a), and that I be added as inventor as follows:

Eran Aharonson, a citizen of Israel, residing at 27 Hasaifan St., Ramat Hasharon, Israel.

I hereby state that the error in inventorship occurred without deceptive intent on my part. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Respectfully submitted,

Eran Aharonson

Attorney Docket No.: **1326-US**

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below under my name.

I believe I am the original, first and sole inventor (if only one name is listed below), or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES

the Specification of which

☐ is attached hereto
☒ was filed on **December 7, 2001**
as Application Serial No. **10/005,314**
and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified Specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, 1.56(a).

PRIORITY CLAIM - NON-U.S. OR PCT PATENT APPLICATION(S)

I hereby claim priority benefits under 35 U.S.C. 119(a)-(d) or 35 U.S.C. 365(b) of any foreign application(s) for patent or inventor's certificate listed below, or 35 U.S.C. 365(a) of any PCT international application listed below which designated at least one country other than the United States of America, and have also listed below to my actual knowledge any foreign application for patent or inventor's certificate, or any PCT international application having a filing date before that of the application on which priority is claimed:

APPLICATION NUMBER	COUNTRY/ CONVENTION	DAY/MONTH/YEAR FILED	PRIORITY CLAIMED

PRIORITY CLAIM - U.S. PROVISIONAL PATENT APPLICATION(S)

I hereby claim priority benefits under 35, U.S.C. §119 of any U.S. Provisional Patent Application listed below that has been filed in the United States in accordance with 35 U.S.C. §119(e), or any U.S. Patent Application listed below that has been converted to a U.S. Provisional Application within one (1) year of its filing date:

APPLICATION NUMBER	DAY/MONTH/YEAR FILED	PRIORITY CLAIMED
60/254,644	7 December 2000	YES

PRIORITY CLAIM - U.S. PATENT APPLICATION(S)

I hereby claim the benefit under 35 U.S.C. §120 of any U.S. Patent Application listed below, and, insofar as the subject matter of each of the claims of this application is not disclosed in any prior U.S. Patent Application in the manner provided by the first paragraph of 35 U.S.C. §112, I acknowledge the duty to disclose material information as defined in 37 CFR §1.56(a), which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION NUMBER	DAY/MONTH/YEAR FILED	STATUS - PATENTED, PENDING, ABANDONED

DESIGNATION OF ATTORNEY OR AGENT

I hereby appoint Daniel J. Swirsky (Agent, Registration No. 45,148) and Heidi M. Brun (Agent, Registration No. 34,504) and practitioners at Customer Number 24505 as my/our attorney(s) or agent(s) with full power of substitution and revocation to prosecute the above-identified application and transact all business connected therewith in the United States Patent and Trademark Office.

Please address all correspondence regarding this application to Customer No. 24505, being:

DANIEL J. SWIRSKY
ALPHAPATENT ASSOCIATES LTD.
P.O.B. 2345
BEIT SHEMESH, ISRAEL 99544

Please direct all telephone calls to (US) (516) 620-4572, all facsimiles to (800) 243-2384, and all e-mail messages to dswirsky@alphapatent.com.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

NAME OF INVENTOR: BARUCH, Amit

ADDRESS: 10 Eliahu Hakim Street, Tel Aviv 69120, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR Amit Baruch

DATE 18.05.2005

NAME OF INVENTOR: MOCHARY, Ran

ADDRESS: 16 Shikmim Street, Nes Ziona, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

NAME OF INVENTOR: RIEMER, Itay

ADDRESS: 14 Kakal Street, Givataim 53237, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

NAME OF INVENTOR: BARUCH, Amit

ADDRESS: 10 Eliahu Hakim Street, Tel Aviv 69120, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

NAME OF INVENTOR: MOCHARY, Ran

ADDRESS: 16 Shikmim Street, Nes Ziona, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE 17 March, 2005

NAME OF INVENTOR: RIEMER, Itay

ADDRESS: 26 Kdoshei Shoah St.. Tel-Aviv 69379. ISRAEL

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE 17 March, 2005

NAME OF INVENTOR: BEN-DOR, Nir

ADDRESS: 37/3 Kiriat Shmona Street, Holon 58483 Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE 16/3/05

NAME OF INVENTOR: YADID, Tal

ADDRESS: Hei Beiyar Street, Rosh Haain 48056, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE 16/3/05

NAME OF INVENTOR: AHARONSON, Eran

ADDRESS: 27 Hasaifan Street, Ramat Hasharon 47248, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

NAME OF INVENTOR: BEN-DOR, Nir

ADDRESS: 37/3 Kiriat Shmona Street, Holon 58550, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

NAME OF INVENTOR: YADID, Tal

ADDRESS: Hei Beiyar Street, Rosh Haain 48056, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE _____

NAME OF INVENTOR: AHARONSON, Eran

ADDRESS: 27 Hasaifan Street, Ramat Hasharon 47248, Israel

COUNTRY OF CITIZENSHIP: Israel

SIGNATURE OF INVENTOR _____

DATE 4 May 2..5

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant:	Amit BARUCH, et al.
Serial No.:	10/005,314
Attorney Docket No.:	1326-US
Filed:	December 7, 2001
Title:	VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES
Group Art Unit:	2655
Examiner:	B. Albertalli

REVOCATION/APPOINTMENT OF POWER OF ATTORNEY OR AUTHORIZATION OF AGENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

As a below-named authorized signatory empowered to sign this statement on behalf of Assignee of Record:

ART – ADVANCED RECOGNITION TECHNOLOGIES LTD.

I hereby revoke all previous powers of attorney or authorizations of agent given in the above-identified application and appoint practitioners at Customer Number 24505 and Daniel J. Swirsky (Agent, Registration No. 45,148) and Heidi M. Brun (Agent, Registration No. 34,504) as my/our attorney(s) or agent(s) with full power of substitution and revocation to prosecute the above-identified application and transact all business connected therewith in the United States Patent and Trademark Office.

A statement under 37 CFR 3.73(b) is enclosed.

Please address all correspondence regarding this application to:

CUSTOMER NUMBER 24505

being:

DANIEL J. SWIRSKY
ALPHAPATENT ASSOCIATES LTD.
P.O.B.2345
BEIT SHEMESH, ISRAEL 99544
Please direct all telephone calls to (516) 620-4572, all facsimiles to (206) 374-6672,

and all e-mail correspondence to dswirsky@alphapatent.com.

AUTHORIZED SIGNATORY: James Arnold

TITLE: Chief Financial Officer/Director

SIGNATURE: JR Arnold Jr

DATE: April 11, 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant: Baruch, et al.

Serial No.: 10/005,314
Filed: December 7, 2001

Title: VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES

Group Art Unit: 2655

Examiner: B. Albertall

WRITTEN CONSENT OF ASSIGNEE UNDER 37 CFR § 1.48(a)

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

As a below-named authorized signatory empowered to sign this statement on behalf of Assignee of Record:

ART - ADVANCED RECOGNITION TECHNOLOGIES LTD.

I hereby consent to the addition of the following person as inventor:

Eran Aharonson, a citizen of Israel, residing at 27 Hasaifan St., Ramat Hasharon, Israel.

A statement under 37 CFR 3.73(b) is enclosed.

AUTHORIZED SIGNATORY: James Arnold

TITLE: Chief Financial Officer/Director

SIGNATURE: J R Arnold Jr

DATE: April 11, 2005

PTO/SB/96(08-00)
Approved for use through 10/31/2002 OMB 0651-0031
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

STATEMENT UNDER 37 CFR 3.73(b)

Applicant/Patent Owner: Amit BARUCH,et al.

Application No./Patent No.: 10/005,314 Filed/Issue Date: DECEMBER 7,2001

Entitled: VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES

ART ADVANCED RECOGNITION TECHNOLOGIES LTD. a CORPORATION

(Name of Assignee) (Type of Assignee, e.g., corporation, partnership, university, government agency, etc.)

states that it is:

1. ☒ the assignee of the entire right, title, and interest; or

2. ☐ an assignee of an undivided part interest.
The extent (by percentage) of its ownership interest is _____%.

in the patent application/patent identified above by virtue of either:

A. ☒ An assignment from the inventor(s) of the patent application/patent identified above. The assignment was recorded in the United States Patent and Trademark Office at Reel 012715, Frame 0495, or for which a copy thereof is attached.

OR

B. ☐ A chain of title from the inventor(s), of the patent application/patent identified above, to the current assignee as shown below

1. From:_____ To:_____
The document was recorded in the Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

2. From:_____ To:_____
The document was recorded in the Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

3. From:_____ To:_____
The document was recorded in the Patent and Trademark Office at
Reel _____, Frame _____, or for which a copy thereof is attached.

☐ Additional documents in the chain of title are listed on a supplemental sheet.

☐ Copies of assignments or other documents in the chain of title are attached.
[NOTE: A separate copy (i.e., the original assignment document or a true copy of the original document) must be submitted to Assignment Division in accordance with 37 CFR Part 3, if the assignment is to be recorded in the records of the PTO. See MPEP 302.8]

The undersigned (whose title is supplied below) is empowered to sign this statement on behalf of the assignee.

April 11, 2005
Date

Signature

James Arnold
Typed or printed name

Chief Financial Officer/Director
Title

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Applicant: Baruch, et al.
Serial No.: 10/005,314
Filed: December 7, 2001
Title: VOICE CONTROL SYSTEM WITH MULTIPLE VOICE RECOGNITION ENGINES

Group Art Unit: 2655
Examiner: B. Albertall

INFORMATION DISCLOSURE STATEMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Entry and consideration of the enclosed Form 1449A - Information Disclosure Statement By Applicant is respectfully requested in accordance with 37 C.F.R. §1.97(c)(2), being filed by the Applicant before the mailing date of a final Office Action.

A copy of each document or other information listed on the enclosed Information Disclosure Statement is provided in accordance with 37 C.F.R. §1.98(a)(2).

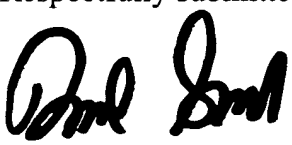
This submission does not represent that a search has been made or that information cited is, or is considered to be, material to patentability as defined in 37 C.F.R. §1.56(b), or that information cited is, or is considered to be "prior art" within the meaning of 35 U.S.C. §§ 102 and 103, or that information cited has been thoroughly reviewed or that any relevance of any portion of a reference is intended. Applicant

reserves the right to establish that any information cited is not "prior art," and that the date of publication indicated for a cited item is in fact different than that which is indicated.

Pursuant to the provisions of M.P.E.P. 609, it is requested that the Examiner return a copy of the enclosed Information Disclosure Statement, marked as being considered and initialed by the Examiner, to the undersigned with the next official communication.

The United States Patent and Trademark Office is hereby authorized to charge Deposit Account No. 501380 for the fee set forth in § 1.17(p) and any other fees that may be required for entry of this paper.

Respectfully submitted,



Daniel J. Swirsky
Representative for Applicant
Registration No. 45,148

ALPHAPATENT ASSOCIATES LTD.
P.O.B. 2345
BEIT SHEMESH, ISRAEL 99544
TEL. (US) 516-620-4573
FAX. (US) 800-243-2384
dswirsky@alphapatent.com

PTO/SB/08A (10-01)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 1 of 2

Complete if Known	
Application Number	10/005,314
Filing Date	December 7, 2001
First Named Inventor	Amit BARUCH,et al.
Group Art Unit	2655
Examiner Name	B. Albertall
Attorney Docket Number	1326-US

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.	U.S. Patent Document Number - Kind Code (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
		US-6,501,966	12-31-2002	Bareis et al.	
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			
		US-			

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.	Foreign Patent Document Country Code	Number	Kind Code (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T

Examiner Signature		Date Considered	

1 Applicant's unique citation designation number (optional). 2 See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. 3 Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). 4 For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. 5 Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. 6 Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

PTO/SB/08B (08-00)
Approved for use through 10/31/2002. OMB 0651-0031
U. S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 2 of 2

Complete if Known	
Application Number	10/005,314
Filing Date	December 7, 2001
First Named Inventor	Amit BARUCH,et al.
Group Art Unit	2655
Examiner Name	B. Albertall
Attorney Docket Number	1326-US

OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T
		Savaraj I. Pawate, Peter Ehlig, "Dialing a Phone by Voice", Machine Design, Jan. 10, 1991	
		Thomas B. Schalk, "Voice Recognition in Cellular Mobile Phones", Speech Technology, Sep./Oct. 1986	

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

1 Unique citation designation number 2 Applicant is to place a check mark here if English language Translation is attached.

AUTOMOTIVE

DIALING A PHONE BY VOICE

SAVARAJ L. PAWATE
PETER EHLIG
hnical Staff Members
exas Instruments Inc.
Houston, TX

Soon you may be able to "dial" a car phone and turn on the lights and wipers with voice commands.

Look for speech recognition to be the next hot technology in the burgeoning automotive electronics industry. In fact, some experts expect voice command systems that control vehicle functions to become widely accepted in this decade.

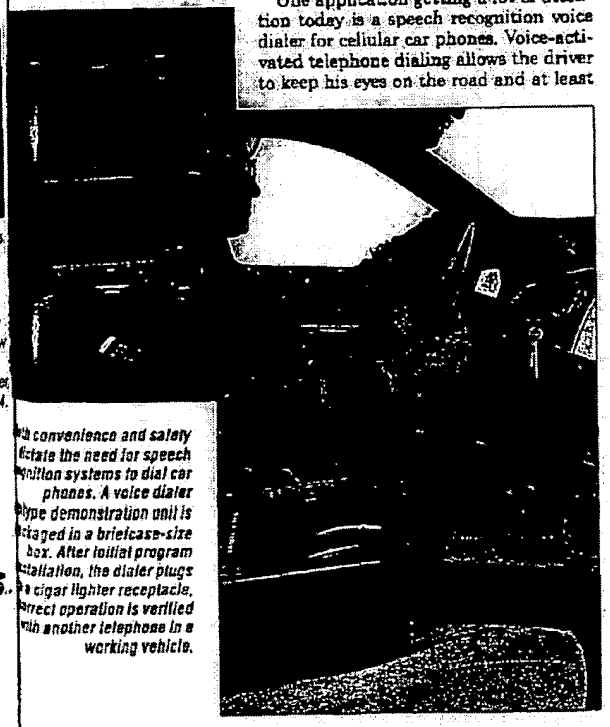
One application getting a lot of attention today is a speech recognition voice dialer for celiular car phones. Voice-activated telephone dialing allows the driver to keep his eyes on the road and at least one hand on the wheel. Conventional dialers, in contrast, require operators to look at a keypad to punch in numbers, a dangerous activity in moving vehicles.

The voice dialer recognizes both male and female voices, as well as a number of dialects. It can have a vocabulary of 25 or more words, depending on memory size. Surprisingly, all this functionality requires only one digital signal processor (DSP).

The voice dialer employs a speech recognition algorithm known as continuous density Hidden Markov Modeling (HMM). HMMs are statistical models for vocabulary words. The algorithms devised to decode voice patterns require substantially more computing power than other techniques, but the improved recognition accuracy outweighs any added expense incurred by using bigger microprocessors.

The voice recognition system has a speaker-independent mode, which means a person does not have to train it to learn his or her voice. For example, any rental-car customer can use the dialer. Any American speakers, regardless of their accents, can be accommodated. Continuous speech recognition is employed so the speaker can talk naturally; no deliberate pauses between words are required.

In addition to unsurpassed accuracy, the voice dialer solves a related communications problem. The cellular telephone industry is rapidly running out of available channels because of the demand for such service. However, a new algorithm called Vector Sum Excited Linear Predictive (VSELP) speech coding, allows the



...h convenience and safety dictate the need for speech ...gnition systems to dial car phones. A voice dialer ...type demonstration unit is ...kaged in a briefcase-size box. After initial program ...stallation, the dialer plugs ...a cigar lighter receptacle. ...rrect operation is verified ...ith another telephone in a working vehicle.

TARGET TOPIC

phone system to accommodate more channels in the available bandwidth than previous methods.

Using the dialer

A typical application uses a grammar definition program built into, or downloaded to, the DSP memory, so either a man or woman can speak to a car telephone and say "Call office" or "Call home." He or she can also state the number to be called, using the words zero through nine for digits or the word "oh" for zero. The user can also define a repertory name, for example, "Call Harvey."

The heart of the dialer comprises fixed-point DSPs, a ROM-based design particularly suited for cellular phones. The DSP has a number of built-in hardware features that speed the implementation of speech recognition algorithms. Consequently, the phones make full use of state-of-the-art digital technology to maximize available telephone channel bandwidth.

Voice dialing features can be added to cellular telephones by simply increasing system memory — other DSP devices are not required. The single speech coding DSP can be time shared to handle voice recognition as well because both functions do not need to run simultaneously. Further, integrated cellular telephones can use the same DSP to control other functions, such as vehicle entertainment equipment, climate, and windshield wipers.

Voice dialer ROM and RAM combinations can be varied to handle different size boot programs, program memory and data. The programs differ depending on the number of telephony applications and the functions provided. An analog interface to the telephone handset, an alpha-numeric display, and interrupt driven connections to the telephone handset complete the set-up.

New product development

To aid in the design of new speech recognition products, the dialer doubles as a development system. An RS-232 interface, for example, supports downloading external software and provides a conduit for control and input information from other systems associated with the dialer. As a result, the voice dialer is easily integrated into a specific application environment or another development system and evaluated.

The RS-232 port downloads to a separate 64k RAM in the voice dialer. The boot program transfers the downloaded program and data to the correct DSP memory.

The dialer has uses other than the telephone application. They include personal computers or workstations where voice

EVERYTHING OLD IS NEW AGAIN

Speech recognition technology is not new. A speaker verification system for military security was introduced in 1974, several years after research began in the 1960s. Even then, the system was said to be superior to fingerprint identification. TI also used a version of the system to control entry to its own computer center.

Today, speech and development systems are designed for a variety of applications, including text-to-speech, record/playback, telephone management, language recognition and speaker verification. Also, credit card verification systems are now widely used.

Text-to-speech algorithms convert ASCII text (as it appears on a computer monitor) into spoken English. The computer-generated voice is natural, intelligible, and has an unlimited vocabulary. Specific applications include inventory assessment, order entry input, and status review.

Record/playback applications are similar to tape recorders or dictation machines. The user can record notes, speeches and other material. However, computer storage provides greater clarity than magnetic recordings and enables the recorded file to be easily merged with other data files.

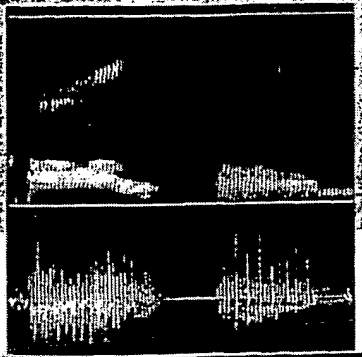
Telephone management systems employ computers to answer telephones, replay messages, and dial other telephones. Applications can be more complex than simple voice mail. In computer banking, for example, customer transactions phoned in can be confirmed at each step of the process by a synthesized voice.

Language recognition enables a computer to recognize complete sentences as they are spoken. One system, for example, can handle applications requiring up to 2,000 words. Language recognition goes beyond mere word recognition; entire sentences are analyzed using context analysis to help determine what is spoken. Language recognition is particularly useful in applications where keyboards cannot be used.

Speaker verification identifies a person through his or her unique voice characteristics. As such, it is ideal for a wide variety of security or entry control applications.

Successful applications arise from a melding of research and development in speech and semiconductor technology, and speech algorithms. For example, established multiple speech databases help create speaker-independent models for the digits used in the voice dialer.

Speech application development requires special software and hardware tools and utilities, and run-time libraries. Such software is available for a variety of DOS and Unix platforms. For example, a speech system tool kit (Speech System V) is available for Xenix or Unix systems running on Intel 80386-based computers. The tool kit also contains an interface for Unix systems operating on minicomputers.



DSP algorithms recognize the digitized form of an analog speech pattern. The top waveshape is a spectrograph of the words call home. The lower waveform is a spectrogram of the same phrase.

recognition is used instead of keyboard input. Also, voice input can supplement factory automation and process inspection data for various machines and computers.

A speech recognition system can also provide hands-off control of a vehicle entertainment system, climate control, windows, windshield wipers, and door locks. For example, a driver can select a radio station with his voice or change the interior temperature without removing his hands from the steering wheel. The voice system can also query the vehicle for fuel status and mpg ratings. Even more elegant features can be had at negligible cost, such as a voice lock that allows the vehicle to be started only by authorized persons.

A demonstration voice dialer system is contained in a portable, briefcase-size box. It is powered by either a 220/110-Vac supply or 12 Vdc through a vehicle cigar lighter receptacle. Such a portable voice dialer can be used as a development system or a test set to diagnose faults in telephones in other mobile units.

The voice dialer circuit is located on one printed-circuit board with programmable array logic (PAL) to minimize the number of individual support logic chips. Voice dialer subsystems include analog circuits and codec, processor and RAM memory, processor control and EEPROMs, display and communications port, and power.

Application-specific grammar

An algorithm can be loaded that makes the dialer recognize up to 25 words without discriminating male or female voices. And application-specific grammar can be either downloaded to the system through an RS-232 serial port, or installed at the factory.

A grammar is also called a sentence model. The DSP and speech recognition algorithms understand and respond to sentence models, and control the syntax by which the words are put together.

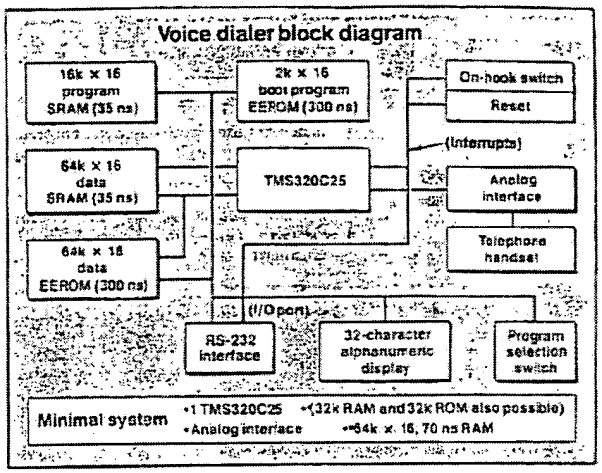
After the grammar is loaded, the voice dialer recognizes the following sequence of commands spoken in any order: call office, call home, or number (digits).

In this sequence, number is a digit string of any length, for example, number 216-666-7777 is a legal sentence. A 1-s pause (or other adjustable value) terminates any speech. When the voice dialer recognizes a complete phrase followed by the pause, it displays a period (.) on the voice dialer 32-character alpha-numeric liquid-crystal display screen. The commands 'enter' or 'cancel' can also terminate the connection.

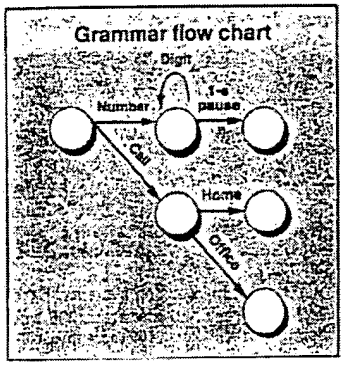
Pressing the off-hook switch on either the voice dialer case or the handset restarts the voice recognition process. In fact, the system recognizes just one command each time the phone goes off hook.

Other application grammars also are possible. An application may, for example, require that the speech recognition system recognize names and the word call as in the command call John Jones.

A basic voice dialer vocabulary consists of 11 digits (zero through nine and the word oh for zero) and four words (call, office, home, and number). But other words are easily added to the application grammar. In one version of the dialer,



The voice dialer requires either a TMS320C25 or TMS320C51 DSP with data memory, program memory, and EEPROM. A telephone handset interface, RS-232 port, display, and various switches comprise a system with a digital configuration that is different for each speech recognition algorithm that it employs.



Flow chart shows operation of the voice dialer when application-specific grammar is loaded. Here, the commands call office, call home, and number (digits) are possible, where digits is a digit string of any length.

TARGET TOPIC

other common words used are enter, cancel, area, code, extension, and emergency.

The database connection

Many speaker-independent word models were created for the voice dialer to eliminate a training phase needed by earlier speech recognition systems. By collecting speech samples from 200 native American speakers (100 male and 100 female), statistical models for each vocabulary word were created. Thus, the likelihood of an unrecognizable word was largely diminished. Care was taken to sample different geographical regions to reflect various dialects. The repertory of voice information is archived in a speech database.

Recognizing that different accents need to be accommodated in certain applications, a speaker-adaptive operating mode was developed. In this mode, the voice dialer boots up with a speaker-independent model. The model is "seed" and the voice dialer controlling algorithm continuously adapts the model to the user in what is called a voice dialer training mode.

Many novel applications also are possible using the database concept. For example, a vocabulary may be developed that is specific to one automobile manufacturer or customer. For some applications, such as a personalized car phone that is disabled when others try to use it, TI could supply speaker-dependent capability for a code word.

In the present voice dialer, all needed voice recognition functions, such as HMM algorithms, signal processing, and grammar control are performed by one DSP. For more complex applications, however, such as large vocabularies and more complex grammars, more than one DSP may be needed. Multiprocessor architecture allows algorithm partitioning so that larger vocabularies may be recognized and accommodated.

Experimental versions of a multiprocessor DSP architecture for speech recognition have already been made. As many as 32 DSPs were connected which, at present, uses an IBM AT computer as a host for development and input/output functions. ■

DSP TARGETED FOR SPEECH RECOGNITION

The newest DSP, the TMS320C51, has an architecture especially configured for speech processing. The design speeds speech-algorithm processing much as a hardware multiplier/accumulator speeds more conventional DSP signal processing.

For example, an important calculation performed by a speech algorithm is selecting a maximum or minimum value out of a set of values. Recognizing this, designers implemented such a maximum and minimum instruction set in hardware for the TMS320C51. A description of the maximum value instruction that compares only two values is illustrated here to help understand the more complex operation for a set of values.

Assume that the maximum value of two numbers is to be found. One is placed in the TMS320C51 accumulator, the other is placed in the accumulator buffer. The instruction CRGT initiates the following sequence: The contents of the accumulator are compared to the contents of the accumulator buffer, and the larger (signed) value is loaded into both registers. A carry bit is modified according to the comparison result. For example, if the contents of the accumulator are greater than or equal to the contents of the accumulator buffer, the carry bit is set to 1; otherwise it is zero. A similar procedure finds the least value in a set of values using the CRLT instruction.

A hardware feature of the TMS320C51 that makes it particularly suited to voice recognition is that, unlike other DSPs, the C51 performs single-cycle 16-×-16-bit multiplications in 35 ns. Data shifting and address manipulations also are in hardware rather than microcode or software.

Speech-recognition algorithms typically are arithmetic intensive and need to access as much DSP power as possible. The C51 DSP features a zero-overhead context switch on interrupts. That means no extra cycle time is needed to save or restore data when an interrupt is received. Because no time cycles are used for data save/restore, that time is available for computation.

The TMS320C51 is a fifth-generation digital signal processor and a fixed-point machine. Available in a 132-pin Quad-Flat-Pack package, the 5-V static CMOS Harvard architecture (separate data and program buses) DSP can be tested using the industry-standard JTAG IEEE P1149.1 boundary scan logic. Capable of more than 28 Mips, the DSP features on-chip ROM, program/data RAM, dual-access data RAM, and memory-security. Also on-chip are wait-state generators, serial ports, a hardware timer, five interrupts, and four external user-maskable interrupts, and 64k I/O ports accessed by sixteen 16-bit address lines.

Texas Instruments Speech System V Toolkit is a software development package used with a 80386-based computer to create speech programs. The tool kit provides the environment to make systems for voice recognition, record-and-play, text-to-speech, and telephone management. An option is also available for speaker verification applications in security products.

Voice Recognition in Cellular Mobile Telephones

A speaker-independent system has been developed which allows "hands-free" telephone use.

Thomas B. Schalk
Director of Technology Development
Voice Control Systems
Dallas, TX

THE ADVENT OF CELLULAR TELEPHONES marked the beginning of high quality mobile telephone service. The first mobile cellular service was offered in Chicago during 1983 and now over 100 major cities offer the service. Accompanying this, has been the proliferation of 500,000 mobile telephone users, all of whom face the challenge of manipulating their phones in a mobile environment.

A typical cellular telephone has approximately 20 keys that correspond to the 10 digits and various control functions. Dialing telephone numbers while driving can be dangerous because the user will typically take his or her eyes off the road to manipulate the telephone keypad. Even after the phone number has been dialed, the user must hold the handset which makes shifting gears and using turn signals, among other things, difficult to do. Hence, many cellular phones have an optional remote microphone that is mounted near the visor, and a speaker located somewhere near the driver. After a call has been placed, the remote microphone and speaker are used in a "hands-free" fashion.

The voice-dialing mobile cellular telephone is one of the most exciting and promising applications of speech recognition in telephony. The use of voice input for dialing can alleviate many of the safety problems associated with cellular telephone systems. A speaker-independent voice recognition system for cellular phones has been developed. This voice control unit is designed to operate optimally in driving vehicles. To develop this system, an extensive voice data collection took place.

In this article the performance requirements for this application will be considered first, then data collection procedures and the functional capabilities of the recognizer will be discussed.

Recognition Performance Requirements

The recognition technology used in the voice control unit for the cellular telephone is speaker-independent and operates on isolated speech. For operating the phone by voice, in place of the key pad, only a small fixed vocabulary is needed. The primary advantage of using speaker-independent technology for the cellular application is that the recognizer need not be trained by each user. Since the noise encountered during driving conditions varies tremendously, it is not practical to expect users to train recognizers properly for the mobile environment. For example, should the user train the system with the engine on or off? Should the blower be set at low, medium, or high? On what road surfaces should the user drive during training, and at what speeds? The speaker-independent reference data used for development of the voice control unit were collected under a wide variety of conditions, which accounts for the robustness of the resultant recognizer in various mobile environmental noises.

It should be noted, though, that a speaker-dependent capability would be useful because it offers vocabulary flexibility. This would allow the user to create custom vocabularies that include people's names to facilitate speed dialing. For example, a person could simply pick up the handset and say "speed-dial Bob Smith" and the preprogrammed phone number for Bob Smith would be dialed automatically.

Commercially available speech recognizers exhibit a wide range of recognition performance. To assess performance, recognition accuracy of the system must be measured. In a crude sense, recognition accuracy refers to the percentage of the time that the recognizer correctly classifies an input utterance. It depends on a number of factors, such as whether the system is speaker-dependent or speaker-independent, whether the system is a discrete or connected word recognizer, the difficulty of the vocabulary, the cost of the system, and the environment in which the system is used. The most stringent performance requirement for the cellular application is reliable digit recognition. Since typical phone numbers are 7 digits long, individual digit accuracy must be very high to make dialing phone numbers by voice practical. Therefore, the ability to detect recognition errors and correct them is critical.

There are three types of errors that a recognizer can make. One, the most obnoxious, is called a substitution error. A substitution occurs when an incorrect word is hypothesized for a valid input utterance. For example, if the active vocabulary for a recognizer includes digits and a "two" is hypothesized when a "nine" was actually spoken, then the recognizer is said to have substituted a two for a nine. In general, substitution error rates must be less than 2 percent for user acceptance. The speaker-independent technology developed for the cellular telephone application achieves this performance goal based on measurements from a large data base of speakers who were "naive" and "inexperienced" recognizer users. The "experi-

Voice Dialing can alleviate many of the safety problems of cellular telephone use.

cised" user tends to incur much lower error rates, making the successful recognition of seven consecutive digits quite likely.

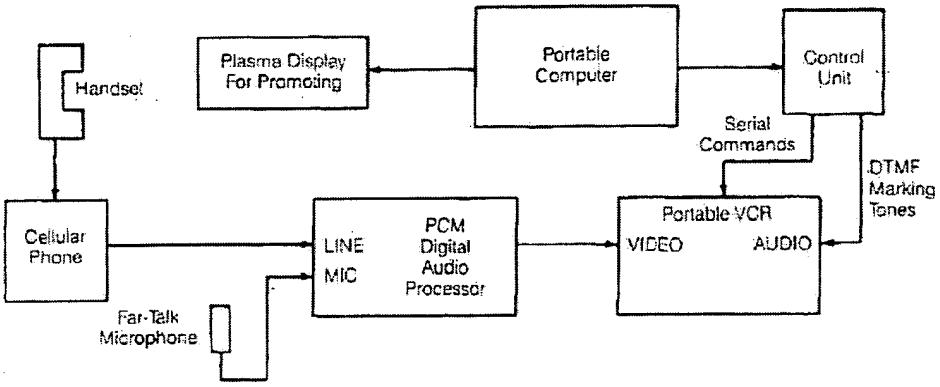
The second type of recognition error is a rejection error. This occurs when a valid input utterance is not classified by the recognizer. When rejections occur, the user simply repeats the utterance—ideally one time—until it is recognized. Rejection errors are not as obnoxious as substitutions, but should not occur more than 5 percent of the time for user acceptance.

The third type of recognition error is a spurious response error. This occurs when an invalid input "sound" (such as a horn honking or uttering a word not found in the vocabulary) is classified as a vocabulary word. Ideally, a recognizer should reject all spurious input. Unfortunately, none of today's recognizers are immune to it. Spurious responses can be minimized by either using push-to-talk microphone arrangements or close-talk microphones. The cellular phone recognition system described here was designed to work under close-talk conditions with a cellular handset. An experimental system using a "far-talk" microphone has also been developed and it uses a push-to-talk microphone activation system to minimize spurious errors.

It is not feasible to quantitatively measure the spurious response error rate for a given recognizer. To do so would involve collecting a data base of all possible sounds that can occur. Nevertheless, if such a data base did exist, spurious response error rates of 50 percent would not be surprising for a typical recognizer. This means that about half the time a spurious sound occurs that is laud enough, the recognizer will attempt to classify it as a word in the vocabulary.

Fig. 1

DATA COLLECTION SYSTEM FOR MOBILE ENVIRONMENT



This data collection process took about 10 minutes per individual.

Data Collection Procedure

Speech recognizers generally exhibit sensitivity to changes in the environments in which they are used. The noise characteristics of moving vehicles vary dramatically depending on the driving conditions and the characteristics of the vehicle itself—say a luxury compared to an inexpensive small car. To develop the speaker-independent reference data for the voice control unit, a large data-base collection was conducted in moving vehicles. Data samples were collected from over 500 people—approximately 100 speakers each in 5 different automobiles selected to span a wide range of driving noise. For safety reasons, the voice donor sat on the passenger side during data collection.

There were two phases to the data collection. The first phase involved the collection of data from a remote ("far-talk") microphone mounted near the visor area. During the second phase, samples were collected through a cellular telephone handset. For each phase of the collection, the voice donor was instructed to say words as they appeared on a custom built prompting display located on the passenger's side of the dashboard. The voice donors were instructed to say the words quickly, in an authoritative manner.

Immediately after the far-talk portion of the collection was completed, the volunteer was again instructed to speak the words as they appeared on the monitor, but this time they were to be said into the cellular handset as though they were "conversing" with it. The entire data collection process took approximately 10 minutes per individual.

The speech samples were recorded using portable digital audio processing recording equipment. This equipment consisted of a SONY Portable VCR SI-2000 and a SONY PCM F-1 Digital Audio Processor (Fig. 1). A small portable computer was used to control the recording equipment and was programmed to feed the speech vocabulary prompts to the prompting display located on the passenger side of the dashboard. In addition, the tapes were automatically marked to indicate speaker number, speech track boundaries, and prompting information. Since only the video portion of the tape was used for PCM speech recording, the normal audio track was available for writing ASCII coded DTMF tones to code this information. This allowed for unsupervised digitization of the speech data onto the VAX computer for speaker-independent vocabulary development. The vocabulary words collected through the handsets included the following list of words:

1. one	15. send
2. two	16. cancel
3. three	17. clear
4. four	18. verify
5. five	19. spouse
6. six	20. home
7. seven	21. friend
8. eight	22. work
9. nine	23. office
10. zero	24. school
11. oh	25. service
12. dial	26. information
13. recall	27. airline
14. speed-dial	28. emergency

During the data-base collection, an emphasis was placed on obtaining a reasonable distribution of different dialects, and on collecting an equal number of male and female voices. As stated earlier, five different cars were used in the collection, as well as a variety of handsets. Most of the speech was collected in a moving vehicle under many different environmental conditions. Some of these conditions were rain with windshield wipers on, defroster on, air conditioner set at various levels, heater set at various levels, windows open, and a ra-

din playing in the background. Furthermore, the data samples were collected in many different locations throughout the Dallas/Ft. Worth Metroplex yielding driving conditions that differed greatly from one location to the next. Prior to the development of the speaker-independent reference data, all of the collected speech was audited and documented. Special attention was given to dialect variations, age of speaker, and any extraordinary characteristics associated with either the speaker or the driving conditions at the time of the collection.

Cellular Telephone Voice Control Unit

The cellular telephone voice control unit discussed here has some noteworthy characteristics. The voice recognition technology is speaker-independent, thus there is no user training required. This specialized recognition system was designed specifically for the noise characteristics of the mobile communications environment. It is a

tware-based recognizer that requires a single general purpose microprocessor (Intel 80186) for implementation. The recognition circuit interfaces to the mobile telephone through the bus that connects the phone control unit (located near the vehicle driver) to the transceiver unit (mounted in the vehicle trunk) as indicated in Fig. 2.

The functional operation of the voice unit centers around syntactically structured voice commands from the user, and voice responses from the voice control unit. The command syntax structure is illustrated in Fig. 3. This simple scheme for voice dialing involves a recognition vocabulary of only 28 words (listed earlier). The output channel of the CODEC in the voice recognizer front-end is used for voice responses to guide the user and provide aural feedback for validation of input. As illustrated in Fig. 2, the voice control unit taps into the voice channel-and-control interface on the control bus of the cellular telephone system. The voice control unit recognizes voice commands given to the phone and then issues appropriate commands to operate the telephone.

Each voice command to the phone is acknowledged by the voice control unit through an aural response. If the command is recognized, then a short beep-tone signifies to the user that the voice control unit has recognized the command and is ready for the next one. If the measured signal-to-noise ratio of a detected utterance is below 20 dB, the utterance is rejected and the voice control unit asks the user to "speak louder." If an utterance has an acceptable signal-to-noise ratio, but is not identified with sufficient confidence, it is rejected, and the response "repeat" is issued to the user, indicating he or she should re-enter the command.

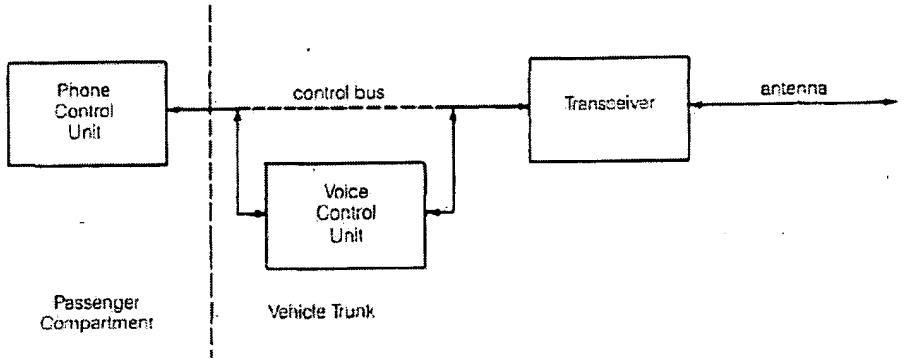
To dial phone numbers, the user simply says "dial" followed by a string of digits. After speaking the last digit of the telephone number, the user says "verify" and a voice response system is activated which repeats the recognized digit sequence through the earpiece of the handset. The user then says "send" in order for that number to be dialed, or "clear" if an improper digit sequence occurred, which activates the top-node vocabulary (dial, recall, speed-dial) and "ready" is synthesized.

For memory dialing (preprogrammed phone numbers), the user says "recall" and then utters a one or two digit sequence, depending on how many preprogrammed numbers the cellular telephone can store. The recognized sequence is repeated and the corresponding phone number in memory is dialed if the word "send" is spoken and recognized. Speed-dialing is acheived by simply picking up the handset and saying "speed dial" followed by one of the ten destination descriptors such as "home," "office," "friend," etc. The recognized words are then repeated and the user can activate the call by saying "send."

Enhanced Cellular Telephone Use

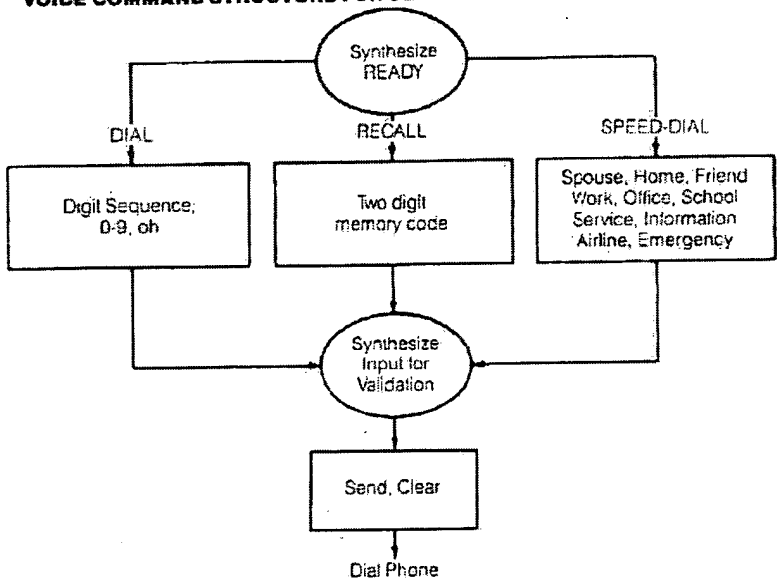
The cellular telephone voice control unit

Fig. 2
VOICE-CONTROLLED MOBILE CELLULAR TELEPHONE



It was designed specifically for noise characteristics of the mobile communications environment.

Fig. 3
VOICE COMMAND STRUCTURE FOR CELLULAR TELEPHONE RECOGNIZER



A 28-word vocabulary is used for this voice-dialing system.

achieves robust speaker-independent speech recognition in a highly variable high-noise environment. Performance measurements using the data base collected in the mobile automotive environment yielded substitution error below 2 percent and rejection rates below 3 percent. The system tested has been implemented in relatively simple hardware which is a fraction of the cost of the cellular telephone unit it enhances. The cellular telephone voice control unit stands as one of the best examples of a fruitful application of speech recognition technology. It achieves its high practical value by simplifying a critical interface and thus significantly enhancing the safety of cellular telephone use.

FOR MORE INFORMATION

Contact Thomas B. Schalk, Voice Control Systems, 14140 Midway Road, Suite 100, Dallas, TX 75244 (214) 386-0301.

THOMAS B. SCHALK *is Director of Technology Development at Voice Control Systems. Currently he directs the VCS research staff, which is working towards advancing speaker-independent speech recognition technology. This effort has led to new technologies that include a phonetic approach for speaker-independent recognition of isolated speech. Prior to joining VCS, he was employed at Texas Instruments, where he conducted speech research and managed a large government contract to develop speech technology.*

Dr. Schalk is experienced in speech data base design and collection procedures and has patents pending for developing speaker-dependent and speaker-independent recognition technology. Dr. Schalk earned a B.S. in Electrical Engineering from George Washington University and received a Ph.D. in Auditory Physiology from the Johns Hopkins School of Medicine.